

## Ein Vergleich von Methoden zur Kurzfristprognose elektrischer Last\*

M. Deistler, W. Fraißler, G. Petritsch und W. Scherrer, Wien

**Übersicht:** Die Prognose des Lastverbrauches ist ein wichtiges Hilfsmittel der Energieversorgungsunternehmen, den Kraftwerkseinsatz zu planen und zu optimieren. Diese Arbeit stellt Prognosemodelle und verschiedene Prognoseverfahren, wie z. B. die Kleinst-Quadrate-Methode, adaptive Schätzverfahren und die Karhunen-Loève-Entwicklung, dar. Diese Verfahren werden auf ihre Prognosequalität hin miteinander verglichen.

### Comparison of methods for short-term load forecasting

**Contents:** Short-term load forecasting is an important instrument for scheduling and optimizing the use of different power-stations. This article describes forecasting-models and forecasting-methods, as the Least-squares method, adaptive methods and the Karhunen-Loève-expansion. The accuracy of the forecasts using these methods is compared.

### 1 Einleitung

Um den optimalen Einsatz von Kraftwerken zu ermöglichen, ist es notwendig, Prognosen über den Verlauf der Nachfrage nach elektrischer Energie während der nächsten Tage zur Verfügung zu haben. Derartige Prognosen bezeichnet man üblicherweise als „kurzfristig“.

Eines der Ziele dieser Arbeit ist es, statistische Methoden zur Automatisierung dieser Prognose vorzustellen und diese miteinander zu vergleichen. Die automatischen Verfahren sollen die händische Prognose ergänzen oder ersetzen.

Die erste umfangreichere Anwendung statistischer Verfahren im Bereich der Nachfrage nach elektrischer Energie stammt von Davies [4], der diese Methoden allerdings nicht auf die Prognose selbst, sondern auf die Analyse des Zusammenhangs zwischen Wettereinflüssen und Lastverlauf anwandte. Reges Interesse

wurde der automatisierten Lastprognose in der ersten Hälfte der 70er Jahre in den USA zuteil. Bedeutende Arbeiten stammen von Farmer und Potton [6], Christiaanse [3], Lijesen und Rosing [12], Gupta und Yamada [9] sowie von Bohlin [1].

Die Schwierigkeiten in der Anwendung statistischer Verfahren bei der Prognose des Lastverlaufs liegen in der komplexen Struktur der Zeitreihe. Der Lastverbrauch zeigt langfristige (trendartige) Variationen, die von ökonomischen, technischen Einflußfaktoren bestimmt werden, Saisonschwankungen (Winter/Sommer), Wochenschwankungen (Abhängigkeit des Verbrauches vom Wochentag, Feiertag) und Tagesschwankungen.

Neben dieser (Kalender-) Zeitabhängigkeit hat aber vor allem das Wetter (Temperatur, Wind, Luftfeuchtigkeit) einen entscheidenden Einfluß auf den Lastverbrauch.

Die geographischen und klimatischen Gegebenheiten des Versorgungsgebietes, die Verbraucherstruktur etc. haben einen starken Einfluß auf die Qualität der Prognosen. Diese Unterschiede erschweren die Vergleichbarkeit von Prognoseergebnissen.

Das Ziel dieser Arbeit ist es, verschiedene Prognosemodelle und statistische Verfahren zur kurzfristigen Lastprognose miteinander zu vergleichen. Die verwendeten Modelle lassen sich in univariate und multivariate Zeitreihenmodelle unterteilen. Neben verschiedenen Regressionsverfahren wird auch die Karhunen-Loève-Entwicklung zur Zerlegung in un beobachtbare Komponenten diskutiert.

### 2 Die verwendeten Prognosemodelle und -verfahren

#### 2.1 Regressionsmodelle

Die nun folgenden Modelle sind detaillierter in Deistler et al. [5] beschrieben.

Mit  $y_{h,d}$  bezeichnen wir die Last zur Tageszeit  $h$

---

\* Diese Arbeit wurde zum Teil im Rahmen des Projektes S32/02 des Fonds zur Förderung der wissenschaftlichen Forschung finanziert.

(gemessen in Stunden oder Halbstunden) am Tag  $d$ .

$$y_{h,d} = \sum_{l=0}^L \sum_{j=0}^J a_{jl}(h, d) y_{h-j,d-l} + \sum_{k=0}^K \sum_{l=0}^L \sum_{j=0}^J b_{jl}^k(h, d) x_{h-j,d-l}^k + u_{h,d} \quad (1)$$

wobei  $h = 1, \dots, H$  (Tageszeit),  $d = 1, \dots, D$  (Tage),  $a_{00}(h, d) = 0$  und  $x_{h-j,d-l}^0 = 1$  für  $j = l = 0$ , ansonsten 0 (Regressionskonstante).

Mit  $x_{h,d}^k$  werden die exogenen Variablen bezeichnet, wobei auch die Zeit und Dummyvariable als exogen betrachtet werden.

Die exogenen Variablen umfassen Polynomfunktionen der Zeit, Jahresdummys (Trendkomponente), harmonische Funktionen der Zeit, Saisondummys (Saisonkomponente), Dummys für die verschiedenen Tagestypen (Wochenkomponente) sowie „echte“ exogene Variable, wie die Wettervariablen (vor allem Temperaturwerte, wie das Tagesmittel, -maximum und -minimum, aber auch transformierte Temperaturwerte).

Man beachte, daß die Koeffizienten  $a_{jl}$  und  $b_{jl}^k$  zeitvariabel definiert sind. Damit wird der sich permanent ändernden Struktur der Zeitreihe Rechnung getragen.

Meist können die autoregressiven Koeffizienten  $a_{jl}$  gleich null gesetzt werden, d. h., die Dynamik der exogenen Einflüsse drückt sich ausschließlich durch die Koeffizienten  $b_{jl}^k$  aus. Diese Modellierung ist vorteilhaft, da zwischen der Last und den beschriebenen „reinen“ Zeitfunktionen, wie z. B. den Tagesartdummies, offensichtlich kein dynamischer Zusammenhang besteht; d. h., für diese exogenen Variablen sind auch die zeitverzögerten Koeffizienten  $b_{jl}^k$  gleich null; würde nun ein ARX-Modell (mit  $a_{jl}$  ungleich null) verwendet, so müßten die Koeffizienten  $b_{jl}^k$  für alle statischen Einflüsse mit den entsprechend zeitverzögerten Werten angesetzt werden, was eine Unzahl von neuen Parametern bedeutet. Es ist also einfacher, die wenigen dynamischen Einflüsse (hier nur die Temperatur) durch ihre eigenen zeitverzögerten Werte zu modellieren.

Zusätzlich zu dieser exogenen Dynamik modellieren wir die Störgröße  $u_{h,d}$  als autoregressiven Prozeß bezüglich des Index  $d$ . Im einfachsten Fall ist  $u_{h,d}$  weißes Rauschen. Die (meist noch stärkere) Korrelation zwischen aufeinanderfolgenden Zeitpunkten innerhalb des Tages kann in der kurzfristigen Prognose (einen oder mehrere Tage im voraus) nicht mehr genutzt werden, wohl aber zur sehr kurzfristigen Prognose (wenige Stunden voraus), die oft in Form einer Korrektur der kurzfristigen Prognose realisiert wird.

Sind die Parameter dieses autoregressiven Störprozesses bekannt oder geschätzt, so kann man ein

Prewhitening vornehmen. Für einen Prozeß erster Ordnung kann dazu eine Cochran-Orcutt-Transformation durchgeführt werden.

Weiter vereinfacht sich die Gl. (1) dadurch, daß die Wettergrößen nur in Tagesschritten berücksichtigt werden. Die übrigen, von der Zeit abgeleiteten exogenen Größen hängen annahmegemäß nicht von der Tageszeit ab.

Der Tagesschwankung wird dadurch Rechnung getragen, daß jeder Tageszeitpunkt getrennt behandelt wird (Längsprognose). Bedingt durch diese Trennung nach Zeitpunkten innerhalb des Tages (Längsprognose) ergeben sich schließlich 24 bzw. 48 Teilmodelle mit Tagesraster ( $h$  fest):

$$y_{h,d} = \sum_{k=0}^K \sum_{l=0}^L b_l^k(h, d) x_{d-l}^k + u_{h,d} \quad (2)$$

Grundsätzlich stehen zwei Alternativen der Modellierung zur Verfügung: eine möglichst vollständige Erfassung aller Einflußgrößen in einem Modell mit konstanten Koeffizienten oder aber ein Modell mit weniger Einflußvariablen, in dem die Variation der abhängigen Variablen auch durch zeitvariable Koeffizienten erklärt wird.

### 2.1.1 Nichtadaptive Schätzung im Regressionsmodell:

Bei Petritsch [13] werden hauptsächlich zeitkonstante Koeffizienten  $b_l^k$  angenommen und diese mittels der gewöhnlichen Kleinstquadratmethode (OLS) geschätzt. In dieser Arbeit wurde aber auch gewichtete Regression verwendet, um eventuellen zeitvariablen Koeffizienten Rechnung zu tragen. (Neben exponentiell abfallenden Gewichten wurde auch saisonale Gewichtung angewandt, was besonders in Übergangszeiträumen Vorteile bringt.) Der Schwerpunkt lag jedoch auf zeitkonstanten Koeffizienten.

Die Modelle stützen sich, wie schon angedeutet, auf längere Beobachtungszeiträume, was die explizite Modellierung der Trend- und Saisonkomponente erforderte. Der Verzicht auf zeitvariable Regressionskoeffizienten bedeutet zwar in gewissem Sinn eine Vereinfachung, dem stehen aber die explizite Modellierung von Trend und Saison gegenüber, wobei die Beobachtungszeiträume flexibel definiert werden können. Speziell werden Vorjahrszeiträume mitverwendet, die der zu prognostizierenden Jahreszeit entsprechen. Dadurch wird die Prognosevarianz reduziert.

Zudem sind folgende methodische Aspekte der genannten Arbeit hervorzuheben:

Automatische Auswahl der signifikanten Variablen mit dem Leap and Bound-Algorithmus von Furnival [8].

Die Option der Verwendung eines Nachbarmodells als Spezialfall des Dummyvariablen-Modells in Judge et al., Kap. 11 [11]: Hier werden für die Prognose jedes einzelnen Zeitpunktes  $h_0$  (z. B. 10 Uhr) auch  $n$  umgebende Zeitpunkte der vergangenen Tage (etwa 9 und 11 Uhr) mitverwendet. Siehe auch Petritsch [13].

### 2.1.2 Schätzung mit dem Kalman-Filter

In Scherrer [14] wurde das Modell Gl. (2) mittels eines adaptiven Verfahrens, basierend auf dem Kalman-Filter, geschätzt. Das dem Kalman-Filter zugrunde liegende Zustandsraum-Modell wurde für unser Problem folgendermaßen spezifiziert:

$$\mathbf{b}(h, d+1) = \mathbf{b}(h, d) + \mathbf{v}_{h,d} \quad (3)$$

$$y_{h,d} = \mathbf{x}'(h, d) \mathbf{b}(h, d) + u_{h,d} \quad (4)$$

$\mathbf{b}(h, d)$  ist der Vektor der Regressionskoeffizienten  $b_i^k(h, d)$  und  $\mathbf{x}(h, d)$  der Vektor aller (auch der verzögerten) exogenen Variablen  $x_{d-l}^k$ .  $\mathbf{v}_{h,d}$  und  $u_{h,d}$  sind miteinander unkorrelierte Störungen, die als weißes Rauschen (in  $d!$ ) angenommen werden. Weiters müssen  $\mathbf{b}(h, d-l)$  für  $l \geq 0$  und  $\mathbf{v}_{h,d}$  sowie  $\mathbf{x}(h, d)$  und  $u_{h,d}$  unkorreliert sein.

Die Beobachtungsgleichung Gl. (4) ist einfach die Gl. (2) in Vektorform geschrieben. Die Zustandsgleichung Gl. (3) beschreibt die zeitliche Änderung des Vektors  $\mathbf{b}(h, d)$  (man beachte, daß  $\mathbf{b}$  hier eine Zufallsvariable ist), der als „Zustand“ des Systems Lastverbrauch interpretiert werden kann ( $\mathbf{b}(h, d)$  beschreibt z. B., wie die Last auf die Temperatur reagiert).

Das Kalman-Filter ergibt eine rekursive Schätzung des Zustandes  $\mathbf{b}(h, d)$  (siehe Jazwinsky [10]).

Es sei  $\mathbf{Q}$  die Varianz-Kovarianzmatrix der Störung  $\mathbf{v}$  und  $R$  die Varianz der Störung  $u$ . Wir nehmen an, daß diese Varianzen nicht zeitabhängig sind und daß  $\mathbf{Q}$  eine Diagonalmatrix ist. Die „Design-Parameter“  $\mathbf{Q}$ ,  $R$  bestimmen die Adaptivität des Filters, das heißt die Schnelligkeit, mit der das Filter auf Änderungen des Zustandes  $\mathbf{b}$  reagiert: Bei großem  $\mathbf{Q}$  und kleinem  $R$  werden die alten Beobachtungen niedergewichtet (das heißt, sie werden schnell „vergessen“), und der Schätzer für  $\mathbf{b}$  kann sich gut den neuen Beobachtungen anpassen.

### 2.1.3 Weitere adaptive Schätzverfahren

Neben dem Kalman-Filter existiert noch eine Reihe weiterer Ansätze, zeitvariable Koeffizienten in Regressionsmodellen zu schätzen. In diese Gruppe fallen Erweiterungen der Methode der exponentiellen Glättung von Brown [2], sowie diejenige adaptive Schätzung, die erstmals von Wheelwright und Makridakis

[15] vorgestellt wurde und die wir hier als WM-Schätzung bezeichnen.

Bei diesen Verfahren wird das Ausmaß der Adaptivität meist nur durch einen einzigen (skalaren) Parameter erfaßt. Das Kalman-Filter hingegen benötigt die Vorgabe der  $k \times k$ -Matrix  $\mathbf{Q}$  (in der allgemeinsten Form bei  $k$  exogenen Variablen). Diese Vereinfachung verringert einerseits die Schwierigkeiten bei der Wahl des optimalen Parameters, andererseits ist die Flexibilität der Verfahren im Vergleich zum Kalman-Filter stark eingeschränkt. Schon aus dieser Überlegung heraus müßten die einfacheren Verfahren im Vergleich etwas schlechter abschneiden. In Fraißler [7] wird versucht, die WM-Schätzung und die exponentielle Glättung als Spezialfälle des Kalman-Filters darzustellen.

Diese Verfahren werden in Fraißler [7] ausführlich beschrieben und so modifiziert und erweitert, daß sie auf das Problem der kurzfristigen Lastprognose anwendbar werden. Eine kurze Beschreibung der Verfahren soll aber auch hier gegeben werden.

Bei der WM-Schätzung wird die Schätzung  $\hat{\mathbf{b}}(h, d)$  für den Parameter  $\mathbf{b}$  zum Zeitpunkt  $(h, d)$  wie auch beim Kalman-Filter durch die Schätzung zum vorangegangenen Zeitpunkt  $(h, d-1)$  und dem daraus resultierenden Prognosefehler ausgedrückt. Beschreibt man das Regressionsmodell durch Gl. (4), dann lautet der Algorithmus der adaptiven Filterung (in der einfachsten Variante)

$$\hat{\mathbf{b}}(h, d) = \hat{\mathbf{b}}(h, d-1) + 2\alpha \mathbf{x}(h, d) [y_{h,d} - \mathbf{x}'(h, d) \hat{\mathbf{b}}(h, d-1)] \quad (5)$$

mit  $0 < \alpha \leq 1$

Bei der exponentiellen Glättung wird die Vorgangsweise von Brown [2], die auf ganz spezielle Klassen von Funktionen beschränkt ist, so verallgemeinert, daß beliebige (auch empirische) unabhängige Variable als Regressoren auftreten dürfen. Das zugrunde liegende Kalkül ist die Minimierung der diskontierten Quadratsumme der Schätzfehler.

$$\sum_{i=0}^D \alpha^i u_{h,d-i}^2 = \sum_{i=0}^D \alpha^i [y_{h,d-i} - \mathbf{x}'(h, d-i) \hat{\mathbf{b}}(h, d)]^2$$

mit  $0 < \alpha \leq 1$  (6)

so daß die exponentielle Glättung auf ein Problem der diskontierten Kleinstquadrat-Schätzung hinausläuft.

## 2.2 Karhunen-Loève-Entwicklung

Die Karhunen-Loève- (KL-) Entwicklung ist eine Entwicklung nach orthonormalen Vektoren  $\Phi_i$ , die nicht a priori feststehen, sondern aus dem konkreten

Datenmaterial gewonnen werden. Wir wenden sie auf die täglichen Lastganglinien  $y_d = (y_{hd})_{h=1, \dots, H}$  oder Teile daraus an. (Siehe auch Deistler et al. [5].)

Zur Erklärung der KL-Entwicklung wird vorläufig angenommen, daß die Ganglinien  $y_d$  in  $d$  identisch verteilt sind und der zugrundeliegende Prozeß von der Form

$$y_{h,d} = \sum_{i=1}^H c_{i,d} \Phi_{i,h} \quad \text{mit } h = 1, \dots, H, d = 1, \dots, D \quad (7)$$

ist, wobei die Entwicklungsvektoren  $\Phi_i$  die (nicht-stochastischen) Eigenvektoren der Gl. (8) sind:

$$\mathbf{Z} \Phi_i = \alpha_i \Phi_i \quad (\text{für } i = 1, \dots, H) \quad (8)$$

mit der 2. Moment-Matrix

$$\mathbf{Z} = (Z_{ij})_{i,j=1, \dots, H}, \quad Z_{ij} = E y_{i,d} y_{j,d} \quad (9)$$

Die Entwicklungskoeffizienten werden aus der Umkehrformel

$$c_{i,d} = \sum_{h=1}^H y_{h,d} \Phi_{i,h} \quad \text{für } i = 1, \dots, H \quad (10)$$

gewonnen. Sie sind daher ebenfalls identisch in  $d$  verteilt. Damit wird für alle Tagesganglinien dasselbe Verteilungsgesetz angenommen — zwischen den einzelnen Lastganglinien sind nur zufällige Unterschiede möglich. In der Praxis wird die Approximationsformel

$$y_{h,d} = \sum_{i=1}^I \hat{c}_{i,d} \hat{\Phi}_{i,h} + \hat{\epsilon}_{h,d} \quad \text{mit } h = 1, \dots, H, \quad I < H \\ d = 1, \dots, D \quad (11)$$

verwendet, wobei die Eigenvektoren aus der Gl. (8) geschätzt werden, in der  $\mathbf{Z}$  durch die empirischen Momente

$$\hat{Z}_{ij} = \sum_{d=1}^D y_{j,d} y_{i,d} / D \quad (12)$$

ersetzt wird und  $\hat{c}_{i,d}$  aus Gl. (10) berechnet wird, wobei ebenfalls  $\Phi_{i,h}$  durch  $\hat{\Phi}_{i,h}$  ersetzt wurde. Der Approximationsfehler  $\hat{\epsilon}_{h,d}$  in Gl. (11) hängt von der Anzahl der Entwicklungsfunktionen  $I$  ab und verschwindet bei  $I = H$ .

Durch die Karhunen-Loève-Entwicklung wird die Last in tageszeitabhängige Faktoren (die Eigenvektoren) und tägliche Faktoren (die Entwicklungskoeffizienten) zerlegt.

Zur Prognose wird das beschriebene Modell noch erweitert. Um  $y_d$  zu prognostizieren, müssen nämlich erst die  $c_{i,d}$  prognostiziert werden. Zu diesem Zweck wird im hierarchischen Karhunen-Loève-Modell Gl. (7) die Annahme identisch verteilter  $c_{i,d}$  fallengelassen und durch die folgende Annahme ersetzt:

$$c_{i,d} = \sum_{k=0}^K \sum_{l=0}^L b_l^k(i, d) x_{d-l}^k + v_{i,d} \quad \text{für } i = 1, \dots, I \quad (13)$$

wobei die  $v_{i,d}$  in  $d$  identisch verteilt sind. Die Beziehung Gl. (7) bleibt weiterhin gültig. Grundsätzlich werden die Entwicklungskoeffizienten  $c_{i,d}$  in diesem Modell durch die gleichen exogenen Variablen wie die Lastwerte im „normalen“ Regressionsmodell Gl. (1) beeinflusst. Da  $I$  im allgemeinen deutlich kleiner ist als  $H$  (meist reichen sechs bis acht Eigenvektoren zur Entwicklung aus), ist die Anzahl der Regressionen kleiner als im Modell Gl. (2). Als zusätzlicher Aufwand gegenüber Gl. (2) müssen die Eigenvektoren berechnet werden, jedoch nicht unbedingt täglich.

Durch die Anwendung dieses Modells werden also alle Zeitpunkte eines Tages oder Tagesabschnittes gemeinsam modelliert.

Konkret wurde die KL-Transformation auf die ungefilterten Lastwerte angewandt, wobei als Störprozeß der Entwicklungskoeffizienten wie beim Regressionsmodell ein AR(1) angenommen und mittels autoregressiver Transformation ausgefiltert wurde.

Alternativ dazu wurden kompliziertere, saisonale Filter zum Prewhitening auf die Last- und Wetterwerte angewandt und anschließend die KL-Entwicklung durchgeführt. Auch die getrennte Entwicklung nach Tagestypen (z. B. Werkzeuge bzw. Wochenenden) wurde versucht, beide Modelle bringen aber keine Verbesserung. Dagegen bewährt sich die Auftrennung des Tages in zwei bis drei Abschnitte, die getrennt behandelt werden.

### 3 Ergebnisse

#### 3.1 Daten und Gütekriterien

Für die Untersuchungen stellte die STEWEAG (Steirische Wasserkraft und Elektrizitäts AG) die halbstündlichen Lastdaten des öffentlichen Bedarfs in der Steiermark (1981 bis 1985) sowie die korrespondierenden Temperatur- und Helligkeitsmessungen aus Graz zur Verfügung.

Verglichen wurden die bedingten Einschnittprognosen (d. h. die täglichen Vorhersagen unter der Annahme perfekter Wettervorhersage) im Zeitraum Anfang November bis Mitte Dezember 1984. Dieser Zeitraum bot sich als für die Modellierung kritischer Bereich an, da die Modelle mit Daten, die zum Teil noch vor diesem Zeitraum liegen, geschätzt werden müßten und im Oktober durch rasche Temperaturänderungen und Einschalten der Heizungen eine Strukturänderung zu erwarten ist. Die Prognosen wurden ex ante in dem Sinn durchgeführt, daß die zu prognostizierenden Variablen nicht zur Parameterschätzung der Prognose verwendet wurden (siehe auch Unterabschnitt 2.1.2). Natürlich wird die Qualität der Prognose unter „realen Bedingungen“, bei der das Wetter selbst

prognostiziert werden muß, schlechter sein. In dem in Abschn. 3.3 angegebenen multivariaten Modell vergrößert sich dadurch die Fehlerstreuung um etwa 0,1 Prozentpunkte (siehe auch Petritsch [13], Kap. 6.1.1). Bei den nicht-adaptiven Verfahren wurden i. a. sieben bis 14 Tage ohne neuerliche Schätzung der Modellparameter prognostiziert.

Als Vergleichsstatistiken wurden vor allem die mittlere absolute Abweichung („MAD“) und die Standardabweichung („STD“) (bezogen auf Mittelwert null) im Prognosezeitraum verwendet, und zwar angewandt sowohl auf die „absoluten“ Prognosefehler  $\hat{y}_{nd} = y_{nd} - \hat{y}_{nd}$  (in MW) als auch auf die relativen Fehler  $\hat{y}_{nd}/y_{nd}$  (in %). Es zeigte sich übrigens, daß MAD-Werte und STD-Werte in unserem Zusammenhang beim Vergleich verschiedener Modelle und Verfahren meist die gleiche Reihenfolge herbeiführten. Das eigentliche Gütekriterium wäre durch die Kosten gegeben, die durch eine Fehlprognose verursacht werden, dies ist allerdings nicht operabel.

### 3.2 Allgemeine Aussagen

Bezüglich des Wettereinflusses kann festgestellt werden, daß die Temperatur den stärksten Einfluß besitzt und zu einer signifikanten Verbesserung der Prognosequalität führt. Für die kurzfristige Lastprognose reicht die Verwendung von Tagesmitteln, -maxima und minima aus. Die Dynamik des Temperatureinflusses ist von großer Bedeutung: die Koeffizienten zeitverzögerter Temperatur sind im Winter bis zu einem Lag von etwa vier Tagen signifikant, im Sommer etwas weniger. Darüber hinaus erweist sich die Verwendung verzögerter Temperaturwerte für die tatsächliche Prognose als vorteilhaft, da z. B. bei Eintagesprognosen die Temperaturwerte vergangener Tage schon exakt vorliegen. Die Stärke des Temperatureinflusses schwankt mit der Jahreszeit (in geringerem Ausmaß auch mit dem Tagestyp) und beträgt im Winter bis zu 1,5% Laständerung pro Grad Celsius.

Nichtlinearität spielt im Winter eine eher untergeordnete Rolle, wird im Sommer aber sehr wichtig (siehe z. B. Petritsch [13], Kap. 3.3.5.2).

Wir haben die Erfahrung gemacht, daß alle Modelle und Verfahren im Sommer und Winter etwa gleich große „absolute“ Prognosefehler und daher, aufgrund des höheren Niveaus, im Winter kleinere relative Prognosefehler ergaben.

Das schwerwiegendste Problem beim Temperatureinfluß lag in der unzureichenden Zahl von Meßstellen. Zur Erhöhung der Prognosequalität wären mehrere Meßreihen unbedingt erforderlich (sogar für die Temperaturreihe des weit entfernt liegenden Salzburg wurden signifikante Koeffizienten errechnet!).

Andere Wetterwerte (insbesondere die Helligkeit) konnten nicht wesentlich zur Verbesserung der Lastprognose beitragen. Dies dürfte auch daran liegen, daß sie stärker regional schwanken.

Bei der Modellierung des Trends bewährte sich der Ansatz von Dummyvariablen für die (ein bis zwei) verwendeten Vorjahre, bei der Saison sind die Dummyvariablen für den Sommerzeit-Übergang und spezielle Zeiträume hervorzuheben.

Die Zuordnung der Tagestypen zu ihren Dummies hat mit Sorgfalt zu geschehen, z. B. können Werkzeuge nach einem Feiertag oft wie Montage behandelt werden. Auch Feiertage lassen sich mit den beschriebenen Verfahren prognostizieren (in erster Näherung als Sonntage), das etwas höhere Fehlerniveau ist auf die geringere Zahl relevanter Vergleichstage zurückzuführen. Dabei sind solche Verfahren vorzuziehen, die auf Vergleichszeiträume des Vorjahres etc. zurückgreifen können.

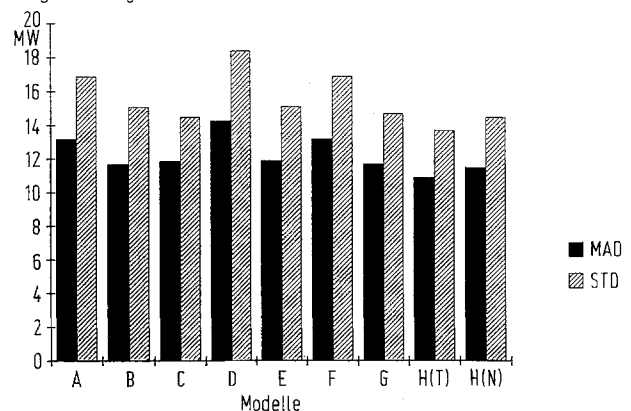
Die meisten der obigen Aussagen lassen sich auch aus anderen untersuchten Datensätzen ableiten (siehe Petritsch [13]).

### 3.3 Nichtadaptive Schätzung im Regressionsmodell

Die hier referierten Ergebnisse stammen von Petritsch [13]; der Prognosezeitraum hierbei war der 7/11/1984 bis 4/12/1984. Genauer untersucht wurden dabei die ersten beiden Wochen (7/11/1984 bis 20/11/1984), über die im folgenden berichtet wird.

Im genannten Zeitraum erzielten univariate Regressionsmodelle (d. h. Modelle ohne Wettervariable) mit exponentiell gewichteter Kleinstquadrate-Schätzung (in Tabelle 1 und Bild 1 mit Variante A bezeich-

Vergleich einiger Modell- und Verfahrensvarianten



**Bild 1.** Vergleich der Ergebnisse einiger Modell- und Verfahrensvarianten (vergleiche auch Tabelle 1); die Bezeichnungen der Modelle und Verfahren entsprechen denen in Tabelle 1; die Abkürzungen H(T) und H(N) entsprechen den Varianten H(Tag) und H(Nacht). Dargestellt werden mittlerer absoluter Fehler und Standardabweichung, jeweils in Megawatt

**Tabelle 1.** Ergebnisse einiger Modell- und Verfahrensvarianten:

| Modell<br>(genauere Erklärung<br>der Modellvarianten<br>im Text)                       | MAD<br>(mittlerer abs.<br>Prognosefehler) |                          | STD<br>(Standardabw. d.<br>Prognf. bezügl. 0) |                          |
|--|---|--------------------------|---|--------------------------|
|  | absol.<br>Fehler<br>in MW                 | relat.<br>Fehler<br>in % | absol.<br>Fehler<br>in MW                     | relat.<br>Fehler<br>in % |
| Verfahren: Nichtadaptive Regressionsschätzer<br>Prognosezeitraum: 7/11/1984–20/11/1984 |   |                          |   |                          |
| A: gew. Regression<br>ohne Temp.var.   | 13,2                                      | 2,4                      | 16,9  | 3,1                      |
| B: OLS<br>mit Temp.var.  | 11,7                                      | 2,2                      | 15,1  | 2,9                      |
| C: gew. Regression<br>mit Temp.var.  | 11,9                                      | 2,2                      | 14,5  | 2,7                      |
| Verfahren: Kalmanfilter<br>Prognosezeitraum: 10/11/1984–20/12/1984                     |   |                          |   |                          |
| D: 6 Dummy   | 14,3                                      | 2,5                      | 18,4  | 3,5                      |
| E: 6 Dummy mit<br>4 Temp.var.  | 11,9                                      | 2,1                      | 15,1  | 2,7                      |
| Verfahren: Karhunen-Loève Entwicklung<br>Prognosezeitraum: 7/11/1984–20/11/1984        |   |                          |   |                          |
| F: ohne Temp.var.  | 13,2                                      | 2,4                      | 16,9  | 3,1                      |
| G: mit Temp.var.   | 11,7                                      | 2,2                      | 14,7  | 2,8                      |
| H: mit T.v.: Tag   | 10,9                                      | 2,0                      | 13,7  | 2,5                      |
| Nacht  | 11,5                                      | 2,2                      | 14,5  | 2,8                      |

net) einen MAD-Wert der relativen Fehler von 2,42%. Dieses Modell wurde mit Daten aus den Zeiträumen 1/10/1982–20/12/1982, 1/10/1983–20/12/1983 und 1/10/1984–6/11/1984 geschätzt. Die unabhängigen Variablen waren zwei Jahresdummies, drei harmonische Schwingungen und fünf Tagestyp-Dummies. Für die Störungen wurde ein autoregressiver Prozeß erster Ordnung angenommen, der Parameter wurde geschätzt oder als 0,5 angenommen. Ohne eine Cochrane-Orcutt-Transformation erhöht sich der MAD-Wert um etwa die Hälfte.

Wie stark die Qualität der Prognose selbst innerhalb dieses einheitlichen Zeitraumes variieren kann, zeigt eine zur Probe durchgeführte Wochen-Prognose (7/11/1984 bis 13/11/1984), die einen MAD-Wert von nur 2,02% ergab.

Unter zusätzlicher Berücksichtigung der Temperatur (konkret der Tagesextrema des aktuellen Tages und von drei Vortagen) wurden mittels OLS-Schätzung (Var. B) ein MAD-Wert von 2,18% und bei exponentiell und saisonal gewichteter Kleinstquadrat-Schätzung (Var. C) ein MAD-Wert von 2,18% erreicht (siehe Tabelle 1 und Bild 1). Auch bei den Modellen mit Temperaturvariablen wurde eine Cho-

crane-Orcutt-Transformation durchgeführt. Ohne eine solche Transformation erhöhen sich die MAD-Werte um ca. 30%.

Die Nutzung von Nachbarinformation bewährte sich nur bei Modellierung über sehr kurze Beobachtungszeiträume, wie sie etwa in Übergangszeiten sowie bei stark schwankenden Datenreihen verwendet wurden.

### 3.4 Adaptive Schätzung im Regressionsmodell

#### 3.4.1 Das Kalman-Filter (siehe Scherrer [14])

Bei der Bestimmung der Diagonalelemente von  $\mathbf{Q}$  (im folgenden mit  $q_{ii}$  bezeichnet) wurde folgendermaßen vorgegangen: Die zu den Tagesdummies bzw. die zu den Temperaturvariablen gehörenden Werte wurden jeweils als gleich angenommen. Schließlich wurden einige Testläufe mit unterschiedlichen Werten für  $q_{ii}$  durchgeführt und dann die Werte ausgewählt, die die minimalen MAD-Werte ergaben. Dabei zeigte sich, daß die Fehlerstatistiken gegenüber der Wahl der  $q_{ii}$  nicht sehr empfindlich reagieren. Dies rechtfertigt auch diese einfache Strategie, die  $q_{ii}$  zu wählen.

Zur Initialisierung des Schätzverfahrens lief die Berechnung vom 1/9/1984 bis zum 9/11/1984. Zur Berechnung der Fehlerstatistiken wurden die Prognosefehler im Zeitraum 10/11/1984 bis 20/12/1984 verwendet.

Ein einfaches univariates Modell ist das „6-Dummy Modell“

$$y_{h,d} = b^0(h, d) + \sum_{k=1,5} b^k(h, d) \delta_{\text{Tagtyp}(d)}^k + u_{h,d}$$

das folgende sechs Tagestypen ( $k = 0, \dots, 5$ ) unterscheidet: „normale“ Arbeitstage, Montage, Freitage, Samstag, Sonn- und Feiertage und Fenstertage.

Wie zu erwarten ist die Basiskomponente  $b^0(h, d)$  „adaptiver“ als die zu den Tagestypen gehörenden Komponenten ( $q_{00} > q_{11} = \dots = q_{55}$ ), da sie ja den saisonbedingten Lastanstieg und den Temperatureinfluß auffangen muß. Es ergab sich ein MAD-Wert von 2,5% (siehe Tabelle 1, Var. D).

Sowohl Pre-Whitening der Lastreihe mit einer Cochrane-Orcutt Transformation als auch „Post-Whitening“, bei dem in einer zweiten Stufe die Prognosefehler mit einem AR-Prozeß modelliert werden, führten zu keiner Verbesserung, da die Autokorrelation der Residuen mit einem Lag von einem bzw. mehreren Tagen sehr klein war.

Es wurde eine Vielzahl von Temperaturmodellen getestet, bei denen zusätzlich zu den sechs Dummies des obigen Modells Temperaturvariable ins Modell aufgenommen werden. Als Beispiel ergab ein Modell mit vier Temperaturvariablen (Mittel und Maximum des aktuellen Tages als auch des Vortages) einen MAD-Wert von 2,1% (siehe Tabelle 1, Var. E).

3.4.2 Vergleich des Kalman-Filters mit anderen adaptiven Verfahren (siehe Fraißler [7])

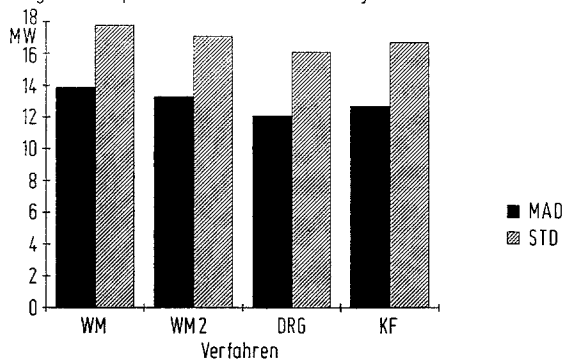
Als einfachste Modelle wurden reine Tagesart-Dummy-Modelle mit vier bzw. sechs Dummy-Variablen untersucht, die sich auf die Modellierung von reinen Tagesarteinflüssen beschränken. Das Sechs-Dummy-Modell ist mit dem oben beschriebenen (Var. D) identisch, beim Vier-Dummy-Modell wurden die Dummies für Montag und Freitag weggelassen und diese Wochentage wie die übrigen Arbeitstage (Dienstag bis Donnerstag) behandelt.

In Sommerperioden sind die Unterschiede zwischen den einzelnen Verfahren gering (siehe Tabelle 2 und Bild 2). In einem Übergangszeitraum (28. Okt. bis 20. Dez. 1984) mit größerem Wettereinfluß und Feier- und Fenstertagen) zeigt sich jedoch bereits eine klare Überlegenheit des Kalman-Filters gegenüber den anderen untersuchten Verfahren (siehe Tabelle 3 und Bild 3). Bezüglich der beiden Verfahrensgruppen läßt sich sagen, daß die WM-Schätzung

**Tabelle 2.** Vergleich einiger adaptiver Verfahren in einem Modell mit vier Dummy-Variablen für Juli 1984

| Verfahren              | MAD<br>(mittlerer abs. Prognosefehler) |                    | STD<br>(Standardabw. d. Prognf. bezügl. 0) |                    |
|------------------------|--|--------------------|--|--------------------|
|                        | absol. Fehler in MW                    | relat. Fehler in % | absol. Fehler in MW                        | relat. Fehler in % |
| WM: Adapt. Filterung   | 13,9                                   | 3,9                | 17,8                                       | 5,3                |
| WMZ: Ad. Flt. (modif.) | 13,3                                   | 3,7                | 17,1                                       | 4,9                |
| DRG: Disk. Regression  | 12,1                                   | 3,3                | 16,1                                       | 4,5                |
| KF: Kalman-Filter      | 12,7                                   | 3,4                | 16,6                                       | 4,6                |

Vergleich adaptiver Verfahren im 4-Dummy Modell für Juli 1984

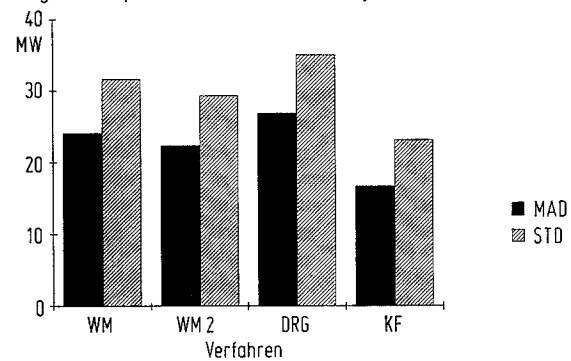


**Bild 2.** Vergleich der Ergebnisse einiger adaptiver Schätzverfahren in einem einfachen 4-Dummy-Modell während eines einfach zu prognostizierenden Sommerzeitraumes. Die Bezeichnungen der Verfahren entsprechen denen in Tabelle 2; dargestellt werden mittlerer absoluter Fehler und Standardabweichung, jeweils in Megawatt

**Tabelle 3.** Vergleich einiger adaptiver Verfahren in einem Modell mit vier Dummy-Variablen für den Übergangszeitraum 28. 10. 1984 bis 22. 12. 1984; aus den Ergebnissen sieht man, daß in diesem Zeitraum der (nicht berücksichtigte) Wettereinfluß von größter Bedeutung ist

| Verfahren         | MAD<br>(mittlerer abs. Prognosefehler) |                    | STD<br>(Standardabw. d. Prognf. bezügl. 0) |                    |
|-------------------|--|--------------------|--|--------------------|
|                   | absol. Fehler in MW                    | relat. Fehler in % | absol. Fehler in MW                        | relat. Fehler in % |
| Adapt. Filterung  | 24,1                                   | 4,5                | 31,7                                       | 6,4                |
| Ad. Flt. (modif.) | 22,4                                   | 4,5                | 29,4                                       | 5,9                |
| Disk. Regression  | 26,9                                   | 5,1                | 35,1                                       | 6,9                |
| Kalman-Filter     | 16,7                                   | 3,1                | 23,3                                       | 4,7                |

Vergleich adaptiver Verfahren im 6-Dummy Modell für Nov./Dez. 1984



**Bild 3.** Vergleich der Ergebnisse einiger adaptiver Schätzverfahren in einem einfachen 6-Dummy-Modell während eines Übergangszeitraumes. Die Bezeichnungen der Verfahren entsprechen denen in Tabelle 2; dargestellt werden mittlerer absoluter Fehler und Standardabweichung, jeweils in Megawatt

im Übergangszeitraum besser abschneidet als die exponentielle Glättung (bzw. diskontierte Kleinstquadrat-Schätzung): sie erfaßt die Variabilität der Koeffizienten besser. Im Sommerzeitraum hingegen liefert die exponentielle Glättung etwas bessere Ergebnisse.

Bei der Hinzunahme von Temperaturvariablen konnten mit den Verfahren der WM-Schätzung keine weiteren Verbesserungen erzielt werden. Im Gegenteil: Für einige Zeiträume zeigten die reinen Dummy-Modelle ein besseres Verhalten als die Modelle, die zusätzlich Temperatureffekte erfassen sollen. Bei der exponentiellen Glättung ergab sich eine Verbesserung, die jedoch geringer war als beim Kalman-Filter.

Gerade in kritischen Zeiträumen, in denen der Wettereinfluß eine große Rolle spielt, zeigt sich die klare Überlegenheit des Kalman-Filters gegenüber den anderen untersuchten (adaptiven) Verfahren.

### 3.5 Karhunen-Loève-Entwicklung

Die Anwendung der Karhunen-Loève-Transformation Petritsch [13] führt in der bedingten Einschnittprognose zu ähnlichen Ergebnissen wie die Regressionsmodelle. Vorteile ergeben sich allerdings in der Mehrschrittprognose und besonders in der „Mittelfristprognose“ (über etwa ein Jahr) und bei der Errechnung von Szenarien.

Die KL-Entwicklung wurde, wie oben beschrieben, zur Transformation von Lastdaten in die „Hauptkomponenten“ verwendet. Dabei erwies sich die gemeinsame Entwicklung über alle Tagestypen am stabilsten, auch Prewhitening vor der Entwicklung führt nicht zu besseren Ergebnissen.

Bei Stundenwerten kann die Entwicklung über den gesamten Tag in einem Abschnitt durchgeführt werden, bei Halbstundendaten bewährt sich die Aufteilung des Tages in zwei bis drei Abschnitte (etwa Tag/Nacht oder Hoch/Niedertarif). Auch wurden überlappende Intervalle verwendet, wobei die Prognosen im Überlappungsbereich einer gewichteten Mittelung unterzogen wurden. Dies brachte jedoch keine erkennbare Verbesserung, was auf eine ausreichend gute Modellierung auch in den Randbereichen der Entwicklungsintervalle schließen läßt.

Im konkreten Vergleichszeitraum 7/11/1984 bis 4/12/1984 erreichte die KL-Transformation, verbunden mit univariaten Regressionsmodellen — vergleichbar der in 3.3 beschriebenen Var. A. —, einen MAD-Wert 2,48% (siehe Tabelle 1, Var. F), also ziemlich identisch mit reinen Regressionsmodellen. Ein multivariates hierarchisches KL-Modell (d. h. mit Berücksichtigung des Temperatureinflusses) — vergleichbar der in 3.3 beschriebenen Var. B — wies  $MAD = 2,18\%$  auf (siehe Tabelle 1, Var. G).

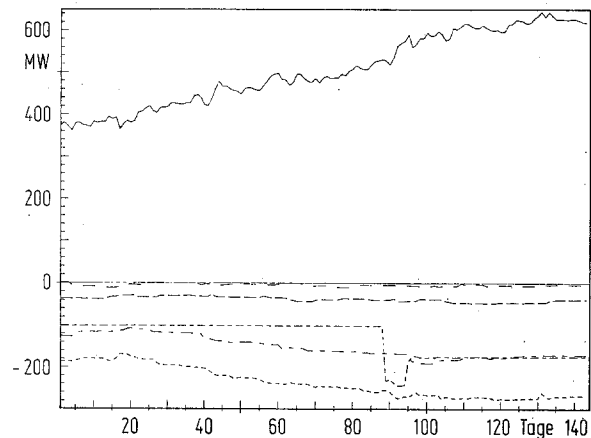
Bei getrennter Entwicklung für den Tages- und Nachtbereich (20 bis 6 Uhr) wurden für den Tag  $MAD = 1,95\%$  und für die Nacht  $MAD = 2,20\%$  (siehe Tabelle 1, Var. H), also etwas bessere Ergebnisse, erzielt.

### 3.6 Vergleich

Vergleiche zwischen verschiedenen Verfahren sind nicht unproblematisch, weil sich z. B. bei adaptiven und nichtadaptiven Verfahren die Datenbasen unterscheiden.

Es zeigt sich, daß die Qualität der hier verglichenen Verfahren keine großen Unterschiede aufweist. Die wesentlichen Unterschiede bestehen in der Art der Modelle und der verwendeten Eingangsinformationen. Eine signifikante Verbesserung kann vermutlich nur durch Verwendung von mehr bzw. genauerer Information (z. B. mehr Wetterdaten, Fernseh-Daten etc.) erzielt werden.

VARIANTE: WS                      ZEIT: 6:00  
SUBVARIANTE: 6-DUMMY  
ALGORITHMUS: KALMAN-FILTER 2 (IMSLIB)



VARIABLEN:

KONST  
DUMMY MONTAG AT  
DUMMY FREITG AT  
DUMMY SAMSTG AT  
DUMMY SONN+FEI  
DUMMY FEN+FRFE.

GRAPHEN:

KONST  
DUMMY MONTAG AT  
DUMMY FREITG AT  
DUMMY SAMSTG AT  
DUMMY SONN+FEI  
DUMMY FEN+FRFE.

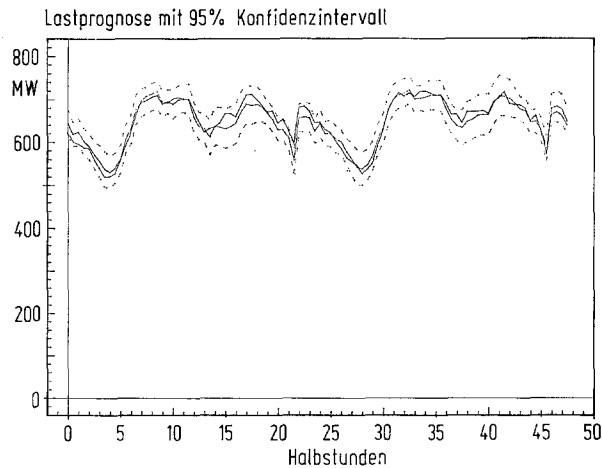
**Bild 4.** Verlauf der Zustände in einem 6-Dummy-Modell bei Verwendung des Kalman-Filters im Zeitraum August 1984 bis Dezember 1984. Man sieht deutlich die saison- und wetterabhängige Veränderung der Zustände (vor allem der ‚Nominallast‘)

Berücksichtigt man neben der Prognosequalität noch weitere Kriterien (Aufwand für die Implementierung, Rechenzeit), so lassen sich die Ergebnisse wie folgt zusammenfassen:

*Adaptive Regressionsschätzer:* Ihre Vorteile liegen in der Effizienz bei einfachen Modellen (mit wenigen Parametern) und entsprechend einfacher Implementierung. Durch ihre geringe Rechenzeit können sie auf jedem Prozeßrechner (zur täglichen Prognose im Lastverteiler) eingesetzt werden. Ihre Nachteile liegen in der schlechten Eignung für die Prognose von Sonder-tagen. Außerdem beeinträchtigen Sondertage auch die Prognosen der folgenden Woche(n). Derzeit noch nicht realisiert sind die adaptive Anpassung der Varianzen  $q_{ii}$  und die Erkennung von Ausreißern. Von den eingesetzten adaptiven Verfahren ist unseres Erachtens dem Kalman-Filter eindeutig der Vorzug zu geben.

*Nicht-adaptive Regressionsschätzer:* Sie eignen sich besonders als Analyseinstrument, zur Errechnung von Szenarien, aber auch für die Prognose von Sonder-tagen oder -zeiträumen, etwa Feiertagsketten, Ur-





**Bild 5.** Vergleich des Verlaufs von Prognose und tatsächlicher Last in einem typischen Zeitraum (12/12/84 bis 13/12/84) bei Verwendung eines Modells mit sechs Dummy-Variablen und aktuellem Tagestemperaturmittel; als Schätzverfahren wurde das Kalman-Filter verwendet

laubsbeginn, Sommerzeitbeginn etc. In der kurzfristigen Prognose besitzen sie etwa dieselbe Qualität wie die adaptiven Schätzer, sind in der Modellierung, der Realisierung auf einem Rechner sowie in der Ausführung aber aufwendiger. Störend wirkt sich auch aus, daß die Verfahrensparameter bislang laufend überwacht werden müssen, was die Verwendung für eine Automatisierung erschwert.

*Karhunen-Loeve-Entwicklung:* Sie eignet sich speziell für längerfristige Prognosen. In der kurzfristigen Prognose ist sie von Qualität und Rechenzeit her mit der Regression zu vergleichen, aber aufwendiger zu realisieren. Auch zur Erkennung falscher Messungen oder von Ausreißern läßt sie sich gut einsetzen.

## Literatur

1. Bohlin, T.: Four case studies of identification of changing systems. In: System Identification — Advances and Case Studies (R. K. Mehra, D. G. Lainiotis (ed.)). New York: Academic Press 1976, ch. 6: Forecasting Load on Power Networks
2. Brown, R. G.: Smoothing, forecasting and prediction of discrete time-series. Englewood cliffs: Prentice-Hall 1962
3. Christiaanse, W. R.: Short-term load forecasting using general exponential smoothing. IEEE Trans. PAS-90 (1971) 900—910
4. Davies, M.: The relationship between weather and electricity demand. The Institution of Electricity Engineers Monograph no. 3148 (1958)
5. Deistler, M.; Fraißler, W.; Petritsch, G.; Scherrer, W.: Methoden zur Kurzfristprognose elektrischer Last.

Institut für Ökonometrie, Operations-Research und Systemtheorie, Abt. für Ökonometrie und Systemtheorie, TU Wien, Institutsarbeit Nr. 82 (1987)

6. Farmer, E. D.; Potton, M. J.: Development of online load-prediction techniques from trials in the south-west region of the CEGB. Proc. IEE 115 (1968) 1549—1558
7. Fraißler, W.: Adaptive Schätzverfahren und ihre Anwendungen in der kurzfristigen Bedarfsprognose elektrischer Energie. Diss. TU Wien, tech. naturw. Fakultät, 1988
8. Furnival, G. M.; Wilson, R. W.: Regressions by leaps and bounds. Technometrics 16 (1974) 499—511
9. Gupta, P. C.; Yamada, K.: Adaptive short-term forecasting of hourly loads using weather information. IEEE Trans. PAS-91 (1972) 2085—2094
10. Jazwinsky, A. H.: Stochastic processes and filtering theory. New York, London: Academic Press 1970
11. Judge, G. E.; Griffiths, W. E.; Hill, R. C.; Lee, T. C.: The theory and practice of econometrics. New York: Wiley 1980
12. Lijesen, D. P.; Rosing, J.: Adaptive forecasting of hourly loads based on load measurements and weather information. IEEE Trans. PAS-90 (1971) 1757—1767
13. Petritsch, G.: ORAKEL: Bedarfsprognose der elektrischen Last mit Regression und Karhunen-Loève Entwicklung. Diss. TU Wien, techn.-naturw. Fakultät, 1986
14. Scherrer, W.: Kurzfristige Lastprognose mit Hilfe des Kalmanfilters. Diplomarbeit an der TU Wien, techn.-naturw. Fakultät, 1986
15. Wheelwright, S. C.; Makridakis, S.: An examination of the use of adaptive filtering in forecasting. Oper. Res. Q. 24 (1973) 55—64

Eingegangen am 14. Januar 1988

Prof. Dipl.-Ing. Dr. M. Deistler

Institut für Ökonometrie, Operations Research und Systemtheorie  
Technische Universität Wien  
Argentinerstraße 8  
A-1040 Wien

Dipl.-Ing. Dr. W. Fraißler

Verbundgesellschaft  
Am Hof 6a  
A-1010 Wien

Dipl.-Ing. Dr. G. Petritsch

EVN (vormals NEWAG-NIOGAS AG)  
Johann Steinböckstraße 1  
A-2344 Maria Enzersdorf am Gebirge

Dipl.-Ing. Mag. W. Scherrer

Institut für Ökonometrie, Operations Research und Systemtheorie  
Technische Universität Wien  
Argentinerstraße 8  
A-1040 Wien  
Österreich